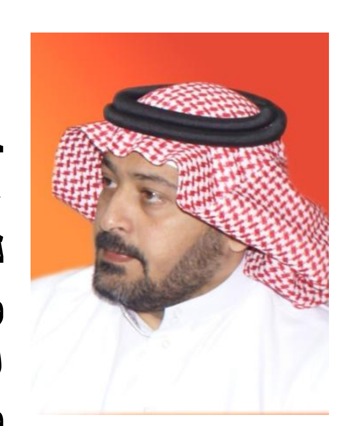

مبادئ علم الإحصاء
إعداد
وليد عبد الرحمن الفرا
رجال ألمع ١٤٢٥هـ

بسم الله الرحمن الرحيم

المحتويات

المقدمة



الحمد لله رب العالمين ، والصلاة والسلام على خاتم النبيين وسيد المرسلين ، نبينا محمد بن عبدالله الهادي البشير الذي بعثه الله رحمة للعالمين ، وعلى آله وأصحابه ، وأنصاره وأتباعه، ومن اهتدى بهديه وعمل بسنته إلى يوم الدين ...

وبعد :

فبحمد الله وعونه جمعت مادة هذا الكتاب المتواضع في فترة وجيزة ووضعت فيه خبرة السنوات الطويلة التي قمت فيها بتدريس مادة الرياضيات للمساهمة في تبسيط مادة الإحصاء لتعدد استخداماتها في شتى المجالات العلمية والعملية، وحاولت قدر طاقتي أن يطابق هذا الكتاب منهج التعليم الثانوي والكليات التي تدرس مبادئ علم الإحصاء.

والله أسأل أن يجعل هذا العمل خالصاً لوجهه الكريم ، وأن يثيبني عليه بقدر مابذلت فيه من جهد ، وأن ينفع به الطلاب والدارسين ؛ كما أهيب بمن يطلع عليه إذا وجد فيه نقصاً أو خطأ أن ينبهني إليه حتى أستدركه في الطبعة القادمة إن شاء الله ، وأن يدعو لي في حياتي وبعد مماتي ، وما توفيقي إلا بالله عليه توكلت وإليه أنيب ، وصلى الله على حبيبنا ونبينا محمد وعلى آله وصحبه أجمعين.

المؤلف

وليد عبد الرحمن الفرا

الفصل الأول

المفاهيم الأساسية لعلم الإحصاء

تمهيد :

الإحصاء ليس غريباً علينا نحن المسلمين، وليس بدعاً بالنسبة لنا الاستفادة من علم الإحصاء في بناء مجتمعنا وتنميته والعمل على سد احتياجاته، وعلى تقدّمه وازدهاره، إذْ أنّ إجراء الإحصاءات وإتباع الأساليب الإحصائية، وجمع المعلومات ، بهدف اتخاذ القرارات بدأت بوقتٍ مبكر منذ بدء بناء المجتمع الإسلامي وتأسيس دولة الإسلام الحنيف.

ومع تطور علوم الرياضيات في القرن الثامن عشر وظهور بعض النظريات العلمية الهامة مثل نظرية الاحتمالات التي كان لها الدور الكبير في تطور هذا العلم واكتسابه أهمية كبرى بحيث أصبح علماً مستقلاً وانتشر استخدامه وبدأ الاهتمام به من قبل العلماء في تطبيق النظريات والطرق والأساليب الإحصائية في الكثير من فروع العلم الحديث كالهندسة والطب والصيدلة والزراعة والصناعة والجغرافيا والفلك وعلم النفس باعتباره الطريقة الصحيحة والأسلوب الأمثل إتباعه في البحث العلمي.

تعريف علم الإحصاء Statistics:

هو ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع وتنظيم وتلخيص وعرض وتحليل البيانات ، وذلك للوصول لنتائج مقبولة وقرارات سليمة على ضوء هذا التحليل.

وكلمة **statistics** مشتقة من كلمة **status** وتعني الدولة باللاتينية، أو كلمة **statistica** بالإيطالية وتعني الدولة أيضاً .

البيانات و المتغيرات:

أولاً: البيانات Data: هي مجموعة من المشاهدات أو الملاحظات التي تؤخذ أثناء دراسة معينة.

أنواع البيانات :

١) بيانات رقمية (كمية) Quantitative Data تنقسم إلى:

أ) بيانات متقطعة Discrete:البيانات المتقطعة محدودة ويمكن عدها، مثل عدد أفراد الأسرة أو عدد الكتب في حقيبتك ..إلخ.

ب) بيانات متصلة Continuous: البيانات المتصلة يمكن قياسها وتأخذ قيماً غير محدودة وغير معدودة، مثل ضغط الدم أو أطوال الطلاب في الفصل ... إلخ .

توضيح : من الممكن أن يكون ضغط الدم للمريض 120 أو 120.1 أو 120.2 أو 120.3 وهكذا تكون الأعداد متصلة (أي بيانات متصلة) . بينما عدد أفراد الأسرة ممكن أن يكون 2 أو 3 أو 4 أو 5 أو 6 ...... أي هذه الأعداد صحيحة (لا يوجد بها كسور عشرية) فتسمى متقطعة (أي بيانات متقطعة أو منفصلة).

٢) بيانات غير رقمية (وصفية) Qualitative Data تنقسم إلى:
أ) مرتبة : مثل (صغير ، وسط ، كبير)
ب) غير مرتبة : مثل (ذكر ، أنثى) أو ألوان العين أو الشعر أو الجنسية .

ثانياً: المتغيرات Variables (Scales):

المتغيرات إما إحصائية أو عشوائية ، فالمتغير الإحصائي يمثل القيم التي تأخذها ظاهرة ما ، في حين أن المتغير العشوائي هو عبارة عن ظاهرة نوعية أو كمية لا يمكن التنبوء بها بشكل مسبق وتقترن بقيم احتمالية .

أنواع المتغيرات :

١) المتغيرات الاسمية Nominal Variables : لها عدد فئات محدد من دون أي معنى كمي لهذه الفئات ، حيث يمكن تصنيف أفراد المجتمع إلى هذه الفئات دون أفضلية لإحداها على الأخرى (مثل متغير الجنس) وفي معظم الأحيان نعطي أرقاماً لتدل على هذه الفئات، فمثلا نرمز للذكر برقم ١ وللأنثى برقم ٢ ولا تدخل هذه الأرقام في العمليات الحسابية.
٢) المتغيرات الترتيبية Ordinal Variables : وهي ذات عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً . ولكن لا يمكن تحديد الفروق بدقة بين قيم الأفراد المختلفة (علي سبيل المثال : كبير ، وسط ، صغير).
٣) المتغيرات الفئوية Interval Variables : هي تلك المتغيرات الكمية التي يمكن إجراء العمليات الحسابية على قيمها دون أن تتأثر المسافة النسبية بين قيمها ، وقيمة الصفر لا تعني عدم توافر تلك الصفة (على سبيل المثال لو حصل طالب على درجة صفر في اختبار القواعد هذا لا يعني أن الطالب لا يعرف شيئاً عن القواعد ، وإذا قلنا أن درجة الحرارة تساوي صفراً فهذا لا يعني عدم وجود درجة حرارة).

٤) المتغيرات النسبية Ratio Variables : هي متغيرات كمية (ليس لها فئات محددة) والصفر في هذا النوع من المتغيرات يمثل عدم توفر الصفة (مثل المتغير الزمني فإذا قلنا أن الزمن يساوي صفر أي لا زمن هناك ، وأيضاً المسافة عندما تساوي صفراً أي لا مسافة موجودة) فلذلك يكثر استخدام هذا المتغير فيزيائياً .

توضيح : البيانات تصبح متغيرات عند إجراء العمليات الإحصائية عليها، فكما لاحظت سابقاً بأن المتغيرات لها نفس أنواع البيانات، والذي ينطبق على البيانات ينطبق على المتغيرات.

مصادر جمع البيانات :

تنقسم مصادر جمع البيانات اللازمة لأي دراسة إحصائية إلى نوعين : مصادر تاريخية و مصادر ميدانية ، وفيما يلي شرح موجز لهما :

أولاً : المصادر التاريخية : وهو ما يؤخذ من السجلات المحفوظة مثل سجلات المواليد والوفيات ، وكذلك البيانات الواردة في رسائل الماجستير والدكتوراه ، أيضاً البيانات التي يتم نشرها من قبل المنظمات الدولية ... إلخ.

ثانياً : مصادر ميدانية : يتم جمع البيانات بطريقة مباشرة عن طريق اتصال الباحث بالوحدة محل الدراسة ، ويتم جمع البيانات ميدانياً بعدة وسائل منها :

١) المقابلة الشخصية : حيث يقوم الباحث بمقابلة أفراد المجتمع المراد دراسته وتوجيه الأسئلة الواردة في البطاقة الإحصائية لكل فرد وتسجيل إجابته، وبهذا يستطيع الباحث أن يحقق أعلى درجات الدقة في جمع البيانات. ولكن من عيوب هذه الطريقة أنها تستغرق جهداً وتكاليف مادية عالية.

٢) الاستمارة الإحصائية (الاستبيان) : حيث يقوم الباحث بتصميم استمارة تشتمل على أسئلة تحقق أهداف البحث ، مع مراعاة شروط كتابة هذه الاستمارة من حيث الدقة والوضوح في عباراتها وأن لا تشتمل على عبارات أو أسئلة مكررة وتكون مرتبة من الأسهل إلى الأصعب .... إلخ.

أسلوب جمع البيانات :

يتم جمع البيانات بأحد الأسلوبين التاليين :

أولاً : الحصر الشامل : حيث يتم جمع البيانات من جميع أفرد ومجتمع الدراسة ، ويستخدم في المجتمعات الصغيرة مثل (مدرسة ، مصنع ...) أو في حالة تباعد الأزمنة بين الإحصائيات مثل التعداد السكاني . وتكمن قوة الحصر الشامل في إعطاء

الباحث صورة حقيقية كاملة عن مجتمع الدراسة ، ومن عيوبه تكاليفه الباهظة ، وطول المدة الزمنية اللازمة لإجرائه .

ثانياً : العينات : يستخدم أسلوب العينات عند دراسة المجتمعات الكبيرة جداً ، ويمكن تعريف العينة على أنها جزء من المجتمع يُختار بطريقة مناسبة ويمثل جميع خصائص ذلك المجتمع بصدق .

أنواع العينات :

١) العينة العشوائية : هي اختيار عينة بحيث تكون فرص ظهور أي من عناصر (مفردات) المجتمع فيها متساوية .
٢) العينة العمدية أو الفرضية : يتم اختيار العينة بطريقة تناسب أهداف البحث بحيث تتوافر في كل عنصر من عناصر العينة شروط محددة يرى الباحث أنها تساعده على الوصول إلى نتائج أفضل في دراسته (على سبيل المثال اختيار الطلاب الأذكياء لتطبيق دراسة عليهم).
٣) العينة الطبقية : وهي العينة التي يتم اختيارها لتشتمل علي خواص المجتمع بالنسب، فمثلاً إذا كان لدينا مجتمع تعليمي عدده ٣٠٠ ، وكانت نسبة الذكور إلي الإناث ٢ : ٣ وأردنا أن نختار عينة من ٥٠ شخصاً ، فلابد أن نختار ٣٠ ذكراً و ٢٠ أنثي .

مفاهيم أساسية:

١) **المجتمع Population:** المجتمع الإحصائي هو عبارة عن جميع الوحدات موضع الدراسة ، سواء كانت هذه الوحدات أفراد أو أشياء أو قياسات .. إلخ ،مثلاً دراسة أعمار مدرسة ما فالمجتمع الإحصائي هنا هو طلاب المدرسة وقت الدراسة ، وقد يكون المجتمع الإحصائي محدوداً، وقد يكون غير محدود .
٢) **العينة Sample:** جزء صغير من المجتمع يلجأ الباحث عادة إلى دراسته، حيث أن العينة تسحب من المجتمع الإحصائي لغرض دراسة صفاته وخصائصه، لذلك يراعى أن تكون هذه العينة عشوائية أي أن تكون العينة ممثلة للمجتمع تمثيلاً صادقاً، ويمكن الحصول على عينة عشوائية باستخدام أسلوب المعاينة العشوائية. فالطبيب يريد معرفة العلة في دم المريض فلا داعي لسحب كل دم المريض بل يأخذ عينة من الدم لفحصها.

٣) <u>المعلمة والإحصائية:</u>

أ) المعلمة : شيء يميز المجتمع ككل، مثل متوسط الدخل الشهري في دولة معينة، أو متوسط طول الطلاب في مدرسة معينة، أو نسبة الأمية في مجتمع ما، أو نسبة المدخنين في دولة ما..... إلخ.

ب) الإحصائية أو الإحصاءة : هي شيء يميز العينة مثل متوسط الدخل الشهري لعينة مكونة من ١٠٠ أسرة في دولة ما، أو متوسط الطول لعينة مكونة من ٣٠ طالب في مدرسة ما ...... وهكذا.

الفصل الثاني
تبويب وعرض البيانات

بعد جمع البيانات سواءً من المصادر التاريخية أو المصادر الميدانية التي تعرضنا لها في الفصل السابق، تكون البيانات خاماً ليست مرتبة أومنظمة (أي بيانات أولية غير مبوبة) وكل قيمة فيها تسمى مفردة أو مشاهدة، فبذلك يصعب دراستها واستخلاص النتائج منها، ولذلك لا بد من تبويبها وترتيبها بطريقة يسهل من خلالها دراستها واستنتاج بعض النتائج منها .

ويعتبر تبويب وعرض البيانات في جداول من الخطوات الأساسية للحصول على المعلومات واستخلاص النتائج المطلوبة من جمع البيانات.

العرض الجدولي للبيانات :

الهدف من العرض الجدولي للبيانات هو إمكانية تحليل البيانات وتطبيق المقاييس الإحصائية المختلفة عليها للحصول على خصائص مجتمع الدراسة. وفيما يلي سنوضح طريقة تكوين الجداول التكرارية لتفريغ البيانات فيها :

أولاً: الجدول التكراري : يتم فيه تنظيم وتلخيص البيانات الوصفية أو الكمية بما يسمى بالتوزيع التكراري Frequency Distribution ، حيث توزع البيانات على شكل فئات ويحدد عدد الأفراد الذين ينتمون لكل فئة ويسمى هذا العدد بتكرار الفئة ويرمز له عادة بالرمز f .

يتكون الجدول التكراري من ثلاثة أعمدة . العمود الأول يكتب فيه الصفة إذا كانت البيانات وصفية أو يكتب فيه الفئة إذا كانت البيانات كمية ، والعمود الثاني للعلامات وهي عبارة عن حزمة تتكون من خمسة خطوط أربعة منها رأسية والخامسة قطرية، والعمود الثالث للتكرار . وفيما يلي مثال للجدولين التكراريين :

الصفات	العلامات	التكرار (عدد الطلاب)
A	卌 I	6
B	卌 III	8
C	卌 卌 卌 I	16
D	卌 卌 卌 卌 II	22
E	卌 III	8
المجموع		60

الجدول التكراري أعلاه يوضح البيانات الوصفية لتقديرات ٦٠ طالباً حيث A ممتاز ، B جيد جداً ، C جيد ،D مقبول ، E ضعيف .

الفئات	العلامات	التكرار (عدد الطلاب)
59 _ 50	卌 I	6
69 _ 60	卌 III	8
79 _ 70	卌 卌 卌 I	16
89 _ 80	卌 卌 卌 卌 II	22
99 _ 90	卌 III	8
المجموع		60

الجدول التكراري أعلاه يوضح البيانات الكمية لدرجات ٦٠ طالباً موزعة على فئات طول الفئة ١٠حيث أقل درجة ٥٠ وأكبر درجة ٩٩ .

الجدولان السابقان يعتبران أول جدولان لتفريغ البيانات، وفي الخطوة الثانية يتم تحسين الجداول بحذف العمود الأوسط (العلامات) فتصبح الجداول بالصورة التالية :

الصفات	التكرار(عدد الطلاب)
A	6
B	8
C	16
D	22
E	8
المجموع	60

الفئات	التكرار(عدد الطلاب)
50 - 59	6
60 - 69	8
70 - 79	16
80 - 89	22
90 - 99	8
المجموع	60

هنا السؤال الذي يطرح نفسه : **ماهي طريقة عمل الفئات المنتظمة للبيانات الكمية ؟**
الغرض من عمل الفئات هو تجميع القيم المتقاربة في مجموعات، بحيث لا يكون عدد الفئات صغيراً فتضيع معالم التوزيع وتفقد كثيراً من التفاصيل، وأيضاً لا يكون عدد الفئات كبيراً فتضيع فائدة التجميع للفئات.
ولتحديد عدد الفئات وطول كل فئة نتبع الخطوات التالية :

١) من الأفضل ترتيب قيم المشاهدات تصاعدياً أو تنازلياً .
٢) إيجاد المدى R ، وهو الفرق بين أكبر قيمة وأصغر قيمة في البيانات.
٣) حساب عدد الفئات M من المعادلة ، أو يمكن تقديرها تبعاً لخبرة الباحث وهي الأكثر شيوعاً.
٤) إيجاد طول الفئة I .

القوانين الرياضية :

R = أكبر قيمة ـ أصغر قيمة

$M = 1 + 3.3 [\log (n)]$ (حيث n عدد المفردات)

$I = \frac{R}{M}$

مثال تطبيقي :

فيما يلي قيم فواتير هاتف جمعت من ١٠٠ مشترك بمحافظة رجال ألمع :

104 154 70 158 160 142 130 112 138 68 66 184 74 120 86 62 86 82
166 98 68 118 92 148 138 78 120 138 166 156 148 102 122 148 136
184 162 154 142 116 166 188 132 156 96 68 100 136 130 128 190
188 170 102 118 168 82 80 76 196 84 110 98 108 120 124 140 176
114 108 92 76 120 194 156 110 132 146 150 128 136 102 174 102
. 92 152 138 128 122 126 100 172 152 198 76 188 96 140 160 190

المطلوب إظهار البيانات في جدول توزيع تكراري مناسب .

الحل : من الملاحظ أن اللبيانات كمية فلا بد أن توزع البيانات في الجدول على شكل فئات ، ولعمل ذلك لا بد أن نتبع الخطوات التالية :

١) من الأفضل أن نرتب البيانات تصاعدياً:

86 86 84 82 82 80 78 76 76 76 74 70 68 68 68 66 62
108 108 104 102 102 102 102 100 100 98 98 96 96 92 92 92
122 122 120 120 120 120 118 118 116 114 112 110 110
138 136 136 136 132 132 130 130 128 128 128 126 124
152 152 150 148 148 148 146 142 142 140 140 138 138 138
168 166 166 166 162 160 160 158 156 156 156 154 154
198 196 194 190 190 188 188 188 184 184 176 174 172 170

٢) حساب قيمة المدى R ، ويتم ذلك باستخراج أقل قيمة فاتورة وهي 62 وأكبر قيمة فاتورة وهي 198 ثم نستحرج المدى :

$$R = 198 - 62 = 136$$

٣) إيجاد عدد الفئات M :

$$M = 1 + 3.3 [\log (100)] = 1 +(3.3 \times 2)= 1+6.6= 7.6 \approx 8$$

٤) إيجاد طول الفئة l :

$$l = R / M = 136 / 8 = 17$$

٥) نختار أصغر قراءة في البيانات لتكون بداية الفئة وهي 62 ويضاف إليها طول الفئة 17 فنحصل بذلك على بداية الفئة الثانية وهي 79 = 17 + 62 ثم نحسب بداية الفئة الثالثة وهي 96 = 17 + 79 وهكذا ..... حتى ننتهي من الفئات الثمانية ونقوم بتصميم الجدول كالتالي :

الفئات	التكرار
62 –	11
79 –	9
96 –	16
113 –	15
130 –	16
147 –	15
164 –	8
181 –	10
المجمـــــــوع	100

ثانياً : الجدول التكراري النسبي Relative Frequency Table : يتكون الجدول من خانتين ، خانة للفئات والخانة الأخرى للتكرار النسبي .
ويتم حساب التكرار النسبي بقسمة التكرار للفئة على المجموع الكلي للتكرار . ويكون مجموع التكرار النسبي مساوياً للواحد الصحيح.

ثالثاً : الجدول التكراري المئوي Percentage Frequency Table : يتكون الجدول من خانتين ، خانة للفئات والخانة الأخرى للتكرار المئوي.
ويتم حساب التكرار المئوي بقسمة التكرار للفئة على المجموع الكلي للتكرار ثم يُضرب الناتج في ١٠٠ . ويكون مجموع التكرار المئوي مساوياً ١٠٠.

الجدول التالي يوضح التكرار ، والتكرار النسبي ، والتكرار المئوي :

الفئات	التكرار	التكرار النسبي	التكرار المئوي
62 –	11	0.11	%11
79 –	9	0.09	%9
96 –	16	0.16	%16
113 –	15	0.15	%15
130 –	16	0.16	%16
147 –	15	0.15	%15
164 –	8	0.08	%8
181 –	10	0.10	%10
المجمـــــــوع	100	1.00	%100

ماهي طريقة حساب مركز الفئة ؟

مركز الفئة Class Mark : ويرمز له بالرمز (X_i) وهو عبارة عن القيمة الوسطية بين الحد الأعلى U والحد الأدنى L للفئة.
ويتم حسابها بالطريقة التالية :

$X_i = (L + U) / 2$

- لو أردنا حساب مركز الفئة الأول في الجدول السابق ، نحدد الحد الأعلى للفئة وهو 79 والحد الأدنى للفئة 62 ثم نجري العملية الحسابية كما يلي :

$X_i = (62 + 79) / 2 = 141 / 2 = 70.5$

الجدول التالي يوضح مراكز الفئات :

الفئات	مراكز الفئات	التكرار	التكرار النسبي	التكرار المئوي
62 –	70.5	11	0.11	%11
79 –	87.5	9	0.09	%9
96 –	104.5	16	0.16	%16
113 –	121.5	15	0.15	%15
130 –	138.5	16	0.16	%16
147 –	155.5	15	0.15	%15
164 –	172.5	8	0.08	%8
181 –	189.5	10	0.10	%10
المجمـــــوع		100	1.00	%100

ما هي الحدود الحقيقية (الفعلية) للفئات؟

عادة ما تكتب البيانات الإحصائية الملخصة والمنظمة مقربة لأقرب وحدة قياس أو أقرب لنصف وحدة قياس .
فإذا كانت البيانات مقربة لأرقام صحيحة فإننا نطرح من الحد الأدنى نصف (0.5) لنحصل على الحد الأدنى الفعلي ، ونضيف نصف (0.5) للحد الأعلى للفئة .
أما إذا كانت البيانات محسوبة لأقرب رقم عشري فإننا نطرح 0.05 من الحد الأدنى ، ونضيف 0.05 للحد الأعلى .

مثال : لدينا درجات 50 طالباً في مادة الرياضيات بحيث كانت أقل درجة 50 وأعلى درجة 97 ، وحددنا عدد الفئات 5 وطول الفترة 10 وجميع البيانات موضحة في الجدول التكراري التالي:

الحدود المقربة للفئات	الحدود الحقيقية للفئات	مراكز الفئات	التكرار	التكرار النسبي	التكرار المئوي
50-59	49.5-59.5	54.5	3	0.06	6
60-69	59.5-69.5	64.5	5	0.10	10
70-79	69.5-79.5	74.5	18	0.36	36
80-89	79.5-89.5	84.5	16	0.32	32
90-99	89.5-99.5	94.5	8	0.16	16
المجموع			50	1.00	100

نلاحظ أن مركز الفئة لا يتأثر بحدود الفئة سواءً كانت حدوداً مقربة أو حدوداً حقيقية .

رابعاً: الجدول التكرارية المتجمعة: يوجد نوعان من الجداول التكرارية المتجمعة وهما الجدول التكراري الصاعد "Cumulative Frequency "Less Than والجدول التكراري الهابط "Cumulative Frequency "Or More .

لو كان مطلوب في الدراسة الإحصائية معرفة عدد الفواتير التي قيمتها أقل من 96 سنجد أنها 20 = 11 + 9 ولو أردنا معرفة عدد الفواتير التي قيمتها أقل من 113 سنجد أنها 36 = 16 + 11 + 9 وهكذا ..... يسمى هذا الجدول بالمتجمع الصاعد .

أما لو كان المطلوب عدد الفواتير التي قيمتها 62 فأكثر سنجد أنها 100 فاتورة، وإذا كان المطلوب عدد الفواتير التي قيمتها 79 فأكثر فسنجد أنها 89 = 11 – 100 وهكذا ... يسمى هذا الجدول بالمتجمع الهابط .

وفيما يلي مثال على الجدولين الصاعد والهابط :

حدود الفئات العليا	التكرار
< 62 (أقل من ٦٢)	0
< 79 (أقل من ٧٩)	11
< 96 (أقل من ٩٦)	20
< 113 (أقل من ١١٣)	36
< 130 (أقل من ١٣٠)	51
< 147 (أقل من ١٤٧)	67
< 164 (أقل من ١٦٤)	82
< 181 (أقل من ١٨١)	90
< 198 (أقل من ١٩٨)	100
الجدول المتجمع الصاعد	

حدود الفئات الدنيا	التكرار
> 62 (٦٢ فأكثر)	100
> 79 (٧٩ فأكثر)	89
> 96 (٩٦ فأكثر)	80
> 113 (١١٣ فأكثر)	64
> 130 (١٣٠ فأكثر)	49
> 147 (١٤٧ فأكثر)	33
> 164 (١٦٤ فأكثر)	18
> 181 (١٨١ فأكثر)	10
> 198 (١٩٨ فأكثر)	0
الجدول المتجمع الهابط	

التمثيل البياني للجداول التكرارية :

تعرفنا فيما سبق على كيفية تنظيم وتلخيص البيانات الإحصائية بواسطة الجداول التكرارية المختلفة ، وفيما يلي سنوضح كيفية تمثيل الجداول التكرارية بيانياً ، ويهدف التمثيل البياني إلى تبسيط عرض البيانات وسهولة دراستها وتحليل بياناتها .

أهم طرق عرض البيانات :

١) الأعمدة البيانية.
٢) المدرج التكراري.
٣)المضلع التكراري.
٤) المنحنى التكراري.
٥) المنحنيات المتجمعة .
٦) القطاعات الدائرية.

أولاً: الأعمدة البيانية Simple Bar Charts: تستخدم غالباً في تمثيل البيانات الوصفية ، كما يمكن استخدامها للمقارنة بين ظاهرة ما في مجتمعين مختلفين وتسمى في هذه الحالة بالأعمدة المزدوجة Multiple Bar Charts.

مثال : ارسم الأعمدة البيانية لتقديرات مادة الرياضيات لعدد 60 طالباً حسب الجدول التكراري التالي :

التقدير	A ممتاز	B جيد جداً	C جيد	D مقبول	E ضعيف	المجموع
التكرار	6	8	16	22	8	60

الحل: نرسم محورين الأفقي للتقديرات ، والرأسي للتكرار . نحدد عرض العمود على المحور الأفقي وليكن 1 سم . ثم نوزع التكرار على المحور الرأسي بأعداد مناسبة ومسافات متساوية . نرسم كل عمود يبعد عن الآخر 1 سم كالشكل المجاور.

مثال : ارسم الأعمدة البيانية المزدوجة لتقديرات مادة الرياضيات لمدرستين مختلفتين بحيث كان عدد الطلاب لكل مدرسة 60 طالباً حسب الجدول التكراري التالي :

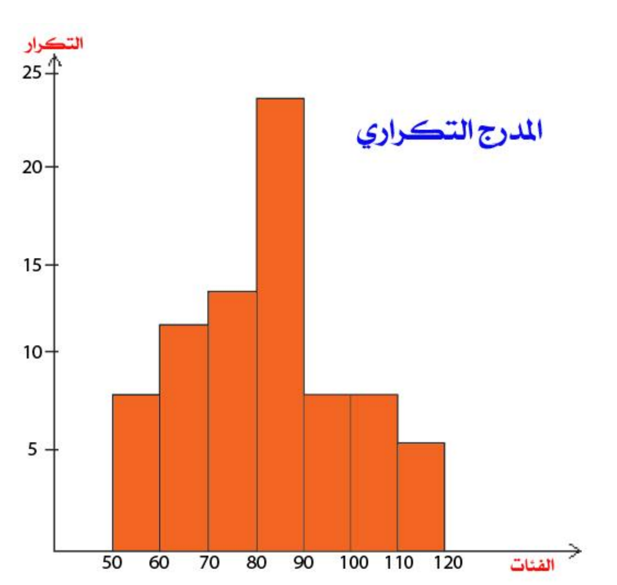
التقدير	A ممتاز	B جيد جداً	C جيد	D مقبول	E ضعيف	المجموع
التكرار للمدرسة أ	6	8	16	22	8	60
التكرار للمدرسة ب	8	10	20	16	6	60

الحل : نستخدم نفس طريقة الأعمدة البيانية في المحاور الرأسية ولكن بالنسبة للمحور الأفقي يقسم مسافات طول المسافة 1سم بحيث في المسافة الأولى نرسم المستطيل الذي يدل على درجة المدرسة أ ، وفي المسافة الثانية نرسم المستطيل الذي يدل على درجة المدرسة ب ، وهكذا .... حتى يصبح الشكل كما هو موضح في الشكل المجاور.

ثانياً: المدرج التكراري: يُرسم المدرَّج التكراري، كما في حالة الأعمدة البيانية، وهو عبارة عن مستطيلات رأسية متلاصقة قاعدة كل منها عبارة عن طول الفئة المناظرة لذلك المستطيل وارتفاع كل منها يساوي (أو يتناسب مع) تكرار تلك الفئة . ويراعى أن يكون تمثيل الفئات على المحور الأفقي حسب حدودها .

مثال : الجدول التالي يمثل الأجر اليومي لثمانين عاملاً في أحد المصانع بالريال حيث وزعنا الأجر على 7 فئات :

الفئات	التكرار
50 –	8
60 –	12
70 –	14
80 –	24
90 –	8
100 –	8
110 –	6
المجموع	80



المدرج التكراري في حالة الفئات الغير منتظمة: إذا كانت الفئات غير متساوية الطول تصبح مساحة المستطيلات الممثلة للمدرج التكراري غير متناسبة مع التكرار وكذلك ارتفاعاتها، لذلك يجب تعديل التكرار قبل رسم المدرج التكراري للفئات غير المتساوية فقط حتى يصبح التكرار المعدل متناسباً مع ارتفاع المستطيل الخاص بالفئة غير منتظمة الطول.

ولتعديل تكرار الفئات الغير منتظمة نقوم بالعملية الحسابية التالية :

$$\text{التكرار المعدل} = \frac{\text{التكرار الفعلي للفئة غير المنتظمة} \times \text{طول الفئة المنتظمة}}{\text{طول الفئة الغير منتظمة}}$$

مثال : الجدول التالي يبين التوزيع التكراري لفئات غير منتظمة لدرجات الطلاب. عدّل التكرارات ثم ارسم المدرج التكراري.

الفئات	50 - 69	70 - 79	80 - 89	90 - 99	المجموع
التكرار	8	18	16	8	50

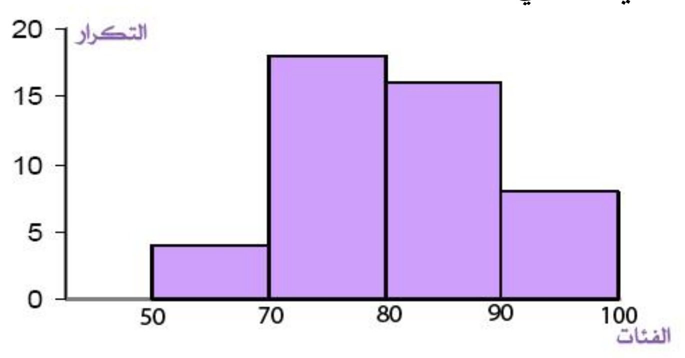
الحل : نلاحظ أن الفئة 50 – 69 غير منتظمة، إذا لا بد من تعديلها كالتالي :

$$\text{التكرار المعدل للفئة } 50 - 69 = \frac{10 \times 8}{20} = 4$$

يصبح الجدول بعد تعديل التكرار كالتالي :

الفئات	50 - 69	70 - 79	80 - 89	90 - 99
التكرار	8	18	16	8
التكرار المعدل	4	18	16	8

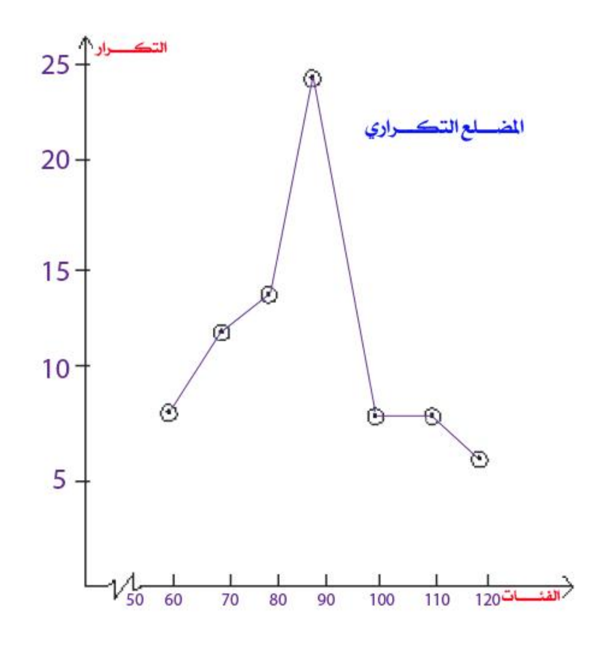
نرسم المدرج التكراري كالتالي :



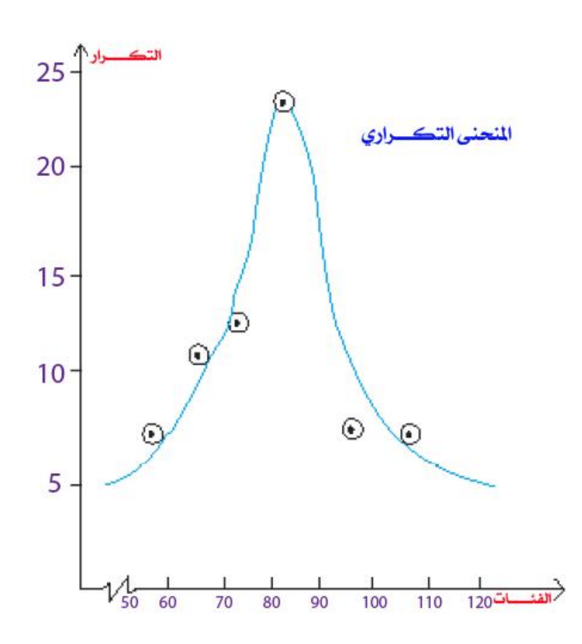
ثانياً: المضلع التكراري: يُرسم المضلع التكراري كما في الحالتين السابقتين من التمثيل البياني، على أن نُحدِّد على المحور الأفقي مراكز الفئات.
تُمثِّل كل فئة من فئات الأجر بنقطة إحداثيها السيني(الأفقي)مركز الفئة وإحداثيها الصادي(الرأسي)التكرار المناظر لتلك الفئة .ثم نوصل هذه النقاط بقطعٍ مستقيمة نحصل على المضلع التكراري.

مثال : الجدول التالي يمثل الأجر اليومي لثمانين عاملاً في أحد المصانع بالريال حيث وزعنا الأجر على 7 فئات وحددنا مراكز الفئات :

الفئات	مراكز الفئات	التكرار
50 –	55	8
60 –	65	12
70 –	75	14
80 –	85	24
90 –	95	8
100 –	105	8
110 –	115	6
المجموع		80



<u>رابعاً: المنحنى التكراري:</u> نتبع في رسم المنحنى التكراري خطوات رسم المضلَّع التكراري نفسها ولكن بدلاً من توصيل كل نقطتين متجاورتين بقطعة مستقيمة بالمسطرة، فإننا نصل كل نقطتين بمنحنٍ ممهَّد باليد أو بشريط مرن ويجب أن يكون المنحني إنسيابياً وحتى لو اضطررنا إلى عدم مروره بعدد قليل من النقاط، بحيث يمر بقربها.

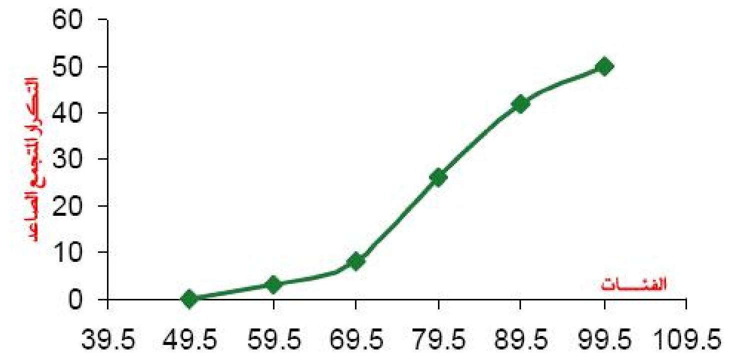


خامساً: المنحنيات المتجمعة:

١) المنحنى المتجمع الصاعد: يرسم على محورين متعامدين الأفقي الفئات والراسي يمثل التكرارات المتجمعة الصاعدة ، ونضع النقاط في الرسم أعلى الحدود الدنيا الحقيقية للفئات بحيث يكون الارتفاع ممثلاً للتكرار المتجمع الصاعد ويمهد المنحنى باليد.

مثال: لدينا جدول تكراري متجمع صاعد به درجات 50 طالباً لمادة الرياضيات بالصورة التالية :

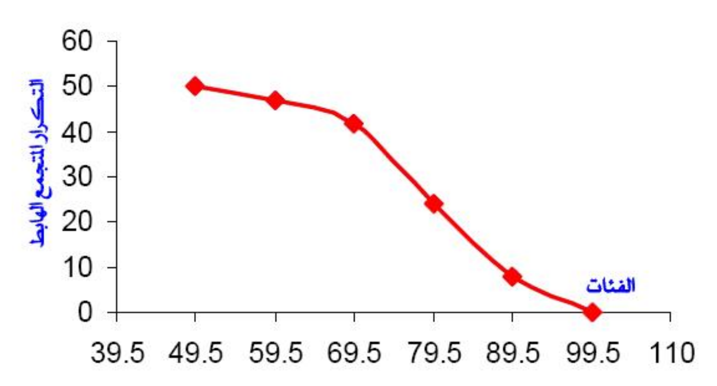
التكرار المتجمع الصاعد	حدود الفئات
0	< 49.5
3	<59.5
8	<69.5
26	<79.5
42	<89.5
50	<99.5



٢) **المنحنى المتجمع الهابط:** يمثل على محورين متعامدين مثل المتجمع الصاعد كما هو موضح في المثال التالي:

مثال: لدينا جدول تكراري متجمع هابط به درجات 50 طالباً لمادة الرياضيات بالتوزيع التالي :

التكرار المتجمع الهابط	حدود الفئات
50	> 49.5
47	>59.5
42	>69.5
24	>79.5
8	>89.5
0	>99.5



سادساً: القطاعات الدائرية Pie Charts: هي عبارة عن دائرة تقسم إلى قطاعات زواياها المركزية تتناسب مع القراءات، ويمكن حساب الزاوية الخاصة بقطاع يمثل قراءة من القراءات كالتالي :

$$\text{زاوية القطاع} = \frac{\text{قيمة (تكرار) الجزء الممثل بالقطاع}}{\text{مجموع القيم (التكرارات)}} \times 360^\circ$$

مثال: البيانات التالية تمثل مؤهلات أعضاء هيئة التدريس في أحد أقسام الجامعة.

المؤهل	دكتوراه	ماجستير	بكالوريوس	دبلوم
العدد	10	16	5	2

مثِّل هذه البيانات بالقطاعات الدائرية .

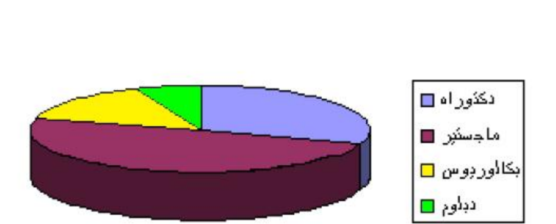
الحل: نرسم الدائرة ثم نحدد الزاوية المركزية لكل قطاع:

١) الدكتوراه $= \frac{10}{33} \times 360^\circ = 109^\circ$

٢) الماجستير $= \frac{16}{33} \times 360^\circ = 175^\circ$

٣) البكالوريوس $= \frac{5}{33} \times 360^\circ = 54^\circ$

٤) الدبلوم $= \frac{2}{33} \times 360^\circ = 22^\circ$



لاحظ أن مجموع الزوايا لا بد أن يساوي 360ْ (109 + 175 + 54 + 22 = 360)

الباب الثالث

المقاييس الإحصائية

رأينا في الباب السابق كيف تم جمع البيانات في جداول تكرارية ثم رسمها بياناً بعدة طرق بهدف الحصول على بعض الخصائص والاتجاهات والعلاقات للمجتمع الإحصائي محل الدراسة، وعادة ما تختلف دقة الرسوم البيانية في نتائجها، وعليه فإنه من الأفضل اللجوء إلى طرق القياس الكمي، حيث يستخدم الباحث الطريقة الرياضية في القياس.

ومن خلال التجارب والملاحظات وجد أن أغلب التوزيعات التكرارية تتمركز القيم عند نقطة متوسطة (قيمة متوسطة) تمثل لمجموعة القيم في التوزيع، ونظراً لأن هذه النقطة أو القيمة تميل للوقوع في المركز داخل مجموعة البيانات فإننا نسميها مقياس النزعة المركزية، وتعتبر القيمة النموذجية الممثلة لمجموعة من البيانات، ويوجد عدة صور لهذه القيمة أهمها وأكثرها شيوعاً (الوسط الحسابي (المتوسط) ، والوسط المرجح ، والوسيط ، والمنوال ، والوسط الهندسي ، والوسط التوافقي) ولكل من هذه المقاييس مزاياه وعيوبه، وفيما يلي سنتناول هذه الطرق بالشرح والتوضيح بأمثلة.

مقاييس النزعة المركزية MEASURES OF CENTRAL TENDENCY

فالهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.

أولاً : الوسط الحسابي (المتوسط) The Arithmetic Mean:

الوسط الحسابي أو المتوسط يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداماً في الإحصاء والحياة العملية إذ يستخدم عادة في الكثير من المقارنات بين الظواهر المختلفة ، ولو أسندت قيمة المتوسط لكل مشاهدة فإن مجموع هذه القيم الجديدة يكون مساوياً لمجموع المشاهدات الأصلية، والوسط الحسابي أو المتوسط ، لمجتمع ما يرمز له بالرمز μ (الحروف اليوناني ميو) ، أما الوسط الحسابي للعينة فيرمز له بالرمز $\overline{X}$ (وتقرأ X bar) .

من ذلك نرى أن الوسط الحسابي يساوي مجموع القيم مقسوماً على عددها. ولحساب الوسط الحسابي نميز بين حالتين هما حالة البيانات غير المبوبة (أي البيانات التي لم يتم وصفها في جداول تكرارية) وحالة البيانات المبوبة (أي التي تم وصفها في جداول تكرارية).

١) <u>البيانات غير المبوبة :</u> إذا كان لدينا n من المشاهدات فالوسط الحسابي يحسب بجمع قيم المشاهدات ثم تُقسم على عددها.

*الوسط للمجموعة n من الأرقام $X_1, X_2, \ldots, X_n$ ويرمز له بالرمز $\overline{X}$ (ويقرأ X bar) ويعرف كالآتي :

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots X_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum X_i}{n}$$

مثال: الوسط الحسابي للأرقام 10،12،5،3، 8 هو:

$$\overline{X} = \frac{8+3+5+12+10}{5} = \frac{38}{5} = 7.6$$

٢) <u>الوسط الحسابى للبيانات المبوبة :</u> إذا كان لدينا عدد k من الفئات ذات المراكز ($X_1, X_2, \ldots, X_k$) ولها تكرارات $f_1, f_2, \ldots, f_k$ على الترتيب فإن الوسط الحسابي يعطى بالعلاقة رقم (١) التالية :

$$\overline{X} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_k x_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{i=1}^{k} f x_i}{\sum_{i=1}^{k} f_i} = \frac{\sum fx}{\sum f} = \frac{\sum fx}{n} \qquad \ldots\ldots\ldots\ldots (١)$$

حيث $n = \sum f$ وهو مجموع التكرارات أي مجموع الحالات.

مثال : إذا كانت 2،6،8،5 تحدث بتكرارات 3,2,4,1 على الترتيب فإن الوسط الحسابي سيكون :

$$\overline{X} = \frac{(3)(5)+(2)(8)+(4)(6)+(1)(2)}{3+2+4+1} = \frac{15+16+24+2}{10} = \frac{57}{10}$$

$$= 5.7$$

مثال : احسب متوسط أعمار الطلاب للبيانات التالية :

فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل: لتبسيط الحل نعمل الجدول التالي :

الفئـــات	مراكز الفئات (x)	التكرار (f)	$x\ f$
5-6	5.5	2	11
7-8	7.5	5	37.5
9-10	9.5	8	76
11-12	11.5	4	46
13-14	13.5	1	13.5
المجموع		20	184

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i = \frac{1}{20}(184) = 9.2 \text{ سنة}$$

ثانياً : الوسط الحسابي المرجح (The Weighted Mean)

في بعض الأحيان نقرن بعض الأرقام $X_1, X_2, \ldots, X_k$ بمعاملات ترجيح أو أوزان $W_1, W_2, \ldots, W_k$ وهذه تعتمد على الدلالة أو الأهمية المرتبطة بهذه الأرقام ، وفي هذه الحالة فإن:

$$\overline{X} = \frac{W_1 X_1 + W_2 X_2 + \ldots + W_k X_k}{W_1 + W_2 + \ldots W_k} = \frac{\Sigma W_i X_i}{\Sigma W_i} = \frac{\Sigma W X}{\Sigma W} \qquad \text{...........} (٢)$$

يسمى بالوسط الحسابي المرجح . ويلاحظ هنا أوجه الشبه بالمعادلة (١) التي يمكن اعتبارها وسطاً حسابياً مرجحاً بأوزان $f_1, f_2, \ldots, f_k$.

مثال : إذا أعطي الامتحان النهائي في مقرر ما وزناً يعادل ثلاثة أمثال الامتحانات الشفوية ، وإذا حصل طالب في الامتحان النهائي على ٨٥ وفي الامتحانات الشفهية على ٩٠ و ٧٠ فإن متوسط تقديره هو:

$$\overline{X} = \frac{(1)(70)+(1)(90)+(3)(85)}{1+1+3}$$

$$= \frac{70+90+255}{5} = \frac{415}{5} = 83$$

خصائص الوسط الحسابي:

فيما يلي سوف نعرض بعض خصائص الوسط الحسابي:

الخاصية الأولى : المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفراً.

$$\sum_{i=1}^{n}(X_i-\overline{X})=0$$

مثال: انحرافات الأرقام 10،12،5،3،8 عن وسطها الحسابي 7.6 هي :

8 – 7.6 = 0.4 , 3 – 7.6 = - 4.6 , 5 – 7.6 = - 2.6 , 12 – 7.6 = 4.4

10 – 7.6 = 2.4

$\Sigma(X_i-\overline{X})$ **= -2.6 - 4.6 + 0.4 + 4.4 + 2.4 = - 7.2 + 7.2 = 0**

الخاصية الثانية : مجموع مربعات انحرافات مجموعة من الأرقام X عن أي رقم a يكون أصغر ما يمكن في حالة واحدة فقط إذا كانت $a=\overline{X}$

الخاصية الثالثة : الوسط الحسابي لعدة مجموعات عبارة عن الوسط الحسابي المرجح لكل وسط حسابي لكل مجموعة مرجحاً بحجم هذه المجموعة.

$$\overline{X}=\frac{f_1m_1+f_2m_2+\dots+f_km_k}{f_1+f_2+\dots+f_k}$$

حيث m_i هي الوسط الحسابي للمجموعة f_i ، i عدد أفراد هذه المجموعة.

مثال: أخذت عينة عشوائية من خمسين عاملاً من عمال أحد مصانع الأثاث فوجد أن متوسط أجر العامل هو ٧٥ ريالاً في اليوم ، ومن عينة أخرى من مائة عامل من عمال أحد مصانع المعلبات فوجد أن متوسط أجر العامل هو ٤٩.٢ ريالاً ، ومن عينة ثالثة لأحد مصانع الحديد والصلب من مائة وخمسين عاملاً وجد أن متوسط أجر العامل اليومي هو ١٠٥ ريال. أوجد الوسط الحسابي لأجر العامل اليومي في المصانع الثلاثة.

الحل :

$$\overline{X} = \frac{50 \times 75 + 100 \times 49.2 + 150 \times 105}{50 + 100 + 150}$$

$$\overline{X} = \frac{3750 + 4920 + 15750}{300} = \frac{24420}{300} = 81.400$$

الخاصية الرابعة : إذا كانت A أي وسط حسابي افتراضي (يمكن أن يكون أي رقم) وإذا كان $d_i = X_i - A$ هو انحرافات X_i عن A فإن :

$$\overline{X} = A + \frac{\sum_{i=1}^{n} d_i}{n} = A + \frac{\Sigma d}{n}$$

$$\overline{X} = A + \frac{\sum_{i=1}^{n} f d_i}{\Sigma f_i} = A + \frac{\Sigma f d}{n}$$

حيث $n = \sum_{j=1}^{n} f_j = \Sigma f$ **، ويمكننا تلخيص المعادلتين السابقتين بالمعادلة :**

$$\overline{X} = A + \overline{d}$$

<u>بعض مميزات الوسط الحسابي :</u>

١) سهولة حسابه .
٢) يأخذ جميع القيم (المشاهدات) في الاعتبار.
٣) أكثر مقاييس النزعة المركزية استخداماً في الإحصاء.

<u>بعض عيوب الوسط الحسابي :</u>

١) يتأثر بالقيم الشاذة المتطرفة (القيم الكبيرة جداً ، والقيم الصغيرة جداً).
٢) يصعب حسابه في حالة الجداول التكرارية المفتوحة لذلك نستخدم الوسيط بدلاً منه في هذه الحالة.
٣) لا يمكن حسابه في البيانات الوصفية.

ثالثاً : الوسيط (The Median)

الوسيط لمجموعة من الأرقام مرتبة حسب قيمها (في ترتيب تصاعدي أو تنازلي) هي القيمة التي تتوسط البيانات التي تقع في المنتصف أو الوسط الحسابي للقيمتين اللتين تتوسطان البيانات أو تقع في منتصف البيانات، أي تقسمها البيانات إلى قسمين متساويين. فإذا كان عدد البيانات فردياً فإن الوسيط يكون المشاهدة التي تقع في المنتصف ، وإذا كان عدد البيانات زوجياً فإن الوسيط هو متوسط المشاهدتين اللتين تقعان في المنتصف .

مثال : **إذا كان لدينا هذه القيم** 165,170,164,172,167,168,173 **التي تمثل أعداد الطلاب في مدارس أحد القطاعات في محافظة رجال ألمع لعام 1424هـ . فأوجد قيمة الوسيط للبيانات السابقة .**

الحل : نرتب البيانات تصاعدياً : 164 , 165 , 167 , 168 , 170 , 172 , 173
نلاحظ أن عدد القراءات السابقة فردية أي 7 = n وبالتالي نطبق القاعدة : (n+1)/2

$$(7 + 1) / 2 = 8 / 2 = 4$$

أي المشاهدة رقم 4 هي الوسيط وتكون : 168 .

مثال : **إذا كان أعداد طلاب الثالث ثانوي طبيعي في مدارس محافظة رجال ألمع كما يلي : 20,21,26,27,29,30,32,33 . أوجد قيمة الوسيط للبيانات السابقة .**

الحل: نرتب البيانات تصاعدياً كما يلي : 20 , 21 , 26 , 27 , 29 , 30 , 32 , 33
نجد أن عدد القراءات 8= n أي زوجية . لا بد هنا أن نأخذ مشاهدتين كالتالي :

$$n/2 = 8/2 = 4$$

$$(n/2) + 1 = (8/2) + 1 = 4 + 1 = 5$$

نأخذ المشاهدتين الرابعة والخامسة ونحسب متوسطهما لاستخراج الوسيط كالتالي:

$$(27 + 29) / 2 = 56 / 2 = 28$$

الوسيط في حالة البيانات المبوبة (الجداول التكرارية) :

لحساب الوسيط حسابياً وبيانياً لا بد أن نتبع الخطوات التالية :

١) نكون الجدول المتجمع الصاعد باستخدام الحدود الحقيقية.
٢) نوجد رتبة الوسيط (n / 2) سواءً كانت n فردية أو زوجية .
٣) نحدد مكان الوسيط بحيث التكرار السابق لـه f_1 والتكرار اللاحق لـه f_2 أكبر من n/2 ، ونأخذ الحد الحقيقي للتكرار السابق على أنه البداية الحقيقية للفئة الوسيطية ونرمز له بالرمز A ، ونعين طول الفئة الوسيطية ويساوي الحد الأدنى للفئة التالية مطروحاً منه الحد الأدنى للفئة الوسيطية ونرمز له بالرمز L ويعطى الوسيط بالعلاقة التالية :

$$\text{Med} = A + \frac{(\frac{n}{2} - f_1)}{f_2 - f_1} L$$

حيث f_1 يمثل التكرار المتجمع الصاعد السابق للتكرار المتجمع الوسيطي ، f_2 يمثل التكرار المتجمع الصاعد اللاحق للتكرار المتجمع الوسيطي .

مثال: احسب الوسيط لأعمار الطلاب المسجلة في الجدول التالي:

فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل : ١) نكون الجدول التكراري المتجمع الصاعد كالتالي :

فئات التكرار المتجمع الصاعد	التكرار المتجمع الصاعد
< 4.5	0
< 6.5	2
< 8.5	7
< 10.5	15
< 12.5	19
< 14.5	20

٢) نحسب (n/2) :

20 / 2 = 10

نجد أن 10 تقع بين 7 , 15 فنضع خطاً أفقياً يمثل تكرار الوسيط المتجمع 10 لنحدد مايلي :

$A = 8.5$, $f_1 = 7$, $f_2 = 15$, $L = 10.5 – 8.5 = 2$

بتطبيق قانون الوسيط نجد أن :

$$Med = 8.5 + \frac{10-7}{15-7} . 2 = 9.25,$$

تحديد الوسيط بيانياً :

يمكننا إيجاد الوسيط بيانياً بثلاث طرق :

من المنحنى المتجمع الصاعد : يتم تحديد النقطة $\frac{n}{2}$ على المحور الرأسي للتكرار ونرسم منها خطاً أفقياً موازياً لمحور الفئات حتى يلتقي بالمنحنى في نقطة ، ومن تلك النقطة نسقط عموداً على المحور الأفقي (الفئات) فتكون نقطة الالتقاء على محور الفئات هي قيمة الوسيط .

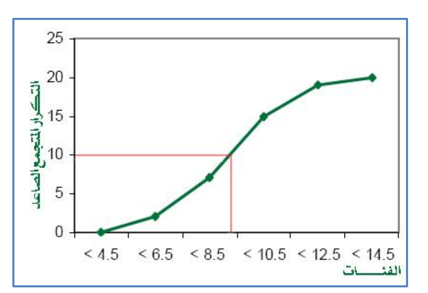
١) من المنحنى المتجمع الهابط : نتبع نفس الخطوات السابقة في المتجمع الصاعد .
٢) من المنحنين المتجمع الصاعد والهابط : نرسم المنحنيين على نفس الرسم ومن نقطة تقاطعهما نسقط عمود على محور الفئات فيدل على قيمة الوسيط.

مثال: احسب الوسيط لأعمار الطلاب المسجلة في الجدول التالي:

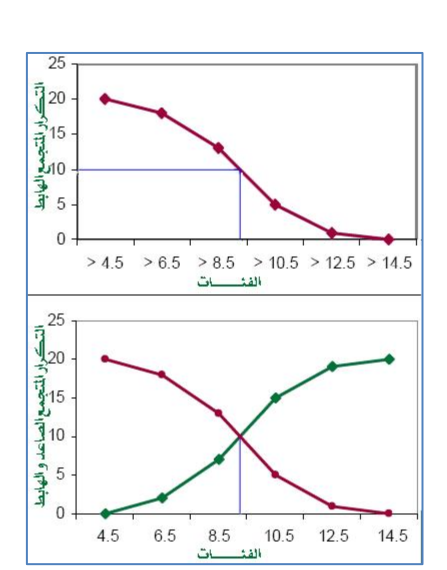
فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل : ١) نكّون الجدول التكراري المتجمع الصاعد والجدول التكراري الهابط كالتالي:

فئات التكرارا المتجمع الهابط	التكرار المتجمع الهابط	فئات التكرار المتجمع الصاعد	التكرار المتجمع الصاعد
> 4.5	20	< 4.5	0
> 6.5	18	< 6.5	2
> 8.5	13	< 8.5	7
> 10.5	5	< 10.5	15
> 12.5	1	< 12.5	19
> 14.5	0	< 14.5	20



٢)نقـوم برسـم المنحنـى المتجمـع الصـاعد ونقـوم بتحديـد النقطـة $\frac{n}{2}$ علـى المحـور الرأسي للتكرار وقيمتها 10 ونرسم منها خطاً أفقيـاً موازيـاً لمحـور الفئـات حتـى يلتقـي بالمنحنى في نقطة ، ومن تلك النقطة نسقط عموداً علـى المحور الأفقي (الفئـات) فتكون نقطـة الالتقـاء علـى محـور الفئـات هـي قيمـة الوسـيط كـم موضـح فـي الرسـم المجاور.



٣) نقوم برسم المنحنى المتجمع الهابط وبنفس الطريقة السابقة نحدد قيمة الوسيط .

٤) نقـوم برسـم المنحنيـين المتجمـع الصـاعد والهابط على نفس الرسم ومن نقطة تقاطعهما نسقط عموداً ليشير لقيمة الوسط على محور الفئات وتساوي 9.25 .

مثال: إذا كان أجر الساعة لخمسة عاملين في مصنع هو ,9.20$, 3.75$, 3.96$ 3.28$, 2.52$ أوجد :

١) وسيط أجر الساعة . ٢) الوسط الحسابي لأجر الساعة .

الحل : ١) بما أن عدد القيم فردي فيكون رتبة المشاهدة = $\frac{n+1}{2} = \frac{5+1}{2} = 3$

المشاهدة الثالثة بعد الترتيب 9.20$, 3.96$, 3.75$, 3.52$, 2.28$ **هي قيمة الوسيط : 3.75**

٢) الوسط الحسابي :

$$\frac{2.52 + 3.96 + 3.28 + 9.20 + 3.75}{5} = \$4.54$$

***لاحظ أن الوسيط لم يتأثر بالقيمة المتطرفة 9.20 بينما تأثر بها الوسط الحسابي . لذلك فإن الوسيط يعطي دلالة أفضل من الوسط الحسابي.**

مثال: يوضح الجدول المقابل أطوال ٤٠ طالباً ، أوجد الطول الوسيط ؟

فئات الطول	التكرار
118- 126	3
127- 135	5
136 - 144	9
145 - 153	12
154 - 162	5
163 - 171	4
172 - 180	2
المجموع	40

الحل: يمكننا حل المثال بطريقتين :

الطريقة الأولى :

الوسيط هو الطول الذي يقع نصف التكرار الكلي قبله $40/2=20$ والنصف الأخر بعده.

بما أن مجموع التكرارات الثلاثة الأولى من الجدول :

$$3 + 5 + 9 = 17$$

ولكي نحصل على الرقم المطلوب 20 لا بد أن نأخذ 3 من تكرار الفئة الرابعة (12) .

وبما أن الفئة الرابعة 145 - 153 هي في الحقيقة تقابل الأطوال 144.5 – 153.5 فإن الوسيط يقع في $\frac{3}{12}$ المسافة بين 144.5 – 153.5 أي أن الوسيط هو :

$$144.5+\tfrac{3}{12}(153.5-144.5)=144.5+\tfrac{3}{12}(9)=146.8Cm$$

الطريقة الثانية :

باستخدام الاستدلال الرياضي كالتالي :

مجموع التكرارات للفئات الثلاث الأولى هي :

$$3 + 5 + 9 = 17$$

مجموع التكرارات للفئات الأربع الأولى هي :

$$3 + 5 + 9 + 12 = 29$$

فإن الوسيط يقع في الفئة الرابعة ، وبالتالي هي الفئة الوسيطية لذلك فإن :

- L_1 الحد الأدنى للفئة الوسيطية — $L_1 = 144.5$
- n مجموع التكرارات — $n = 40$
- $(\Sigma f)_1$ مجموع التكرارات لجميع الفئات قبل الفئة الوسيطية — $(\Sigma f)_1 = 17$
- f median تكرار الفئة الوسيطية — f median = 12
- c طول الفئة الوسيطية — $c = 9$

وبالتعويض نجد أن الوسيط يساوي :

$$L_1+\left(\frac{\frac{n}{2}-(\Sigma f)_1}{f_{median}}\right)C$$

$$=144.5+\left(\frac{\frac{40}{2}-17}{12}\right)(9)=146.8Cm$$

مميزات الوسيط :

١) لا يتأثر بالقيم المتطرفة .
٢) يمكن حساب الوسيط في حالة الجداول التكرارية المفتوحة للبيانات الكمية .
٣) يمكن إيجاده في حالة البيانات الوصفية التي يمكن ترتيبها.

عيوب الوسيط :

١) لا يأخذ جميع القيم في الاعتبار.
٢) لا يسهل التعامل معه في التحاليل الإحصائية والرياضية.

رابعاً: المنوال (The Mode)

يعرف المنوال على أنه القيم الأكثر تكراراً (شيوعاً) في مجموعة البيانات ، وقد يكون لمجموعة البيانات منوالاً واحداً ولذلك تسمى وحيدة المنوال Unimodal ، أو يكون لها منوالين وتسمى ثنائية المنوال Bimodal، أو يكون لها أكثر من منوالين فتسمى متعددة المنوال ، وقد لا يكون لمجموع البيانات أي منوال وبذلك تسمى عديمة المنوال.

مثال: مجموعة البيانات التالية : 22 , 5 , 7 , 9 , 9 , 9 , 10 , 10 , 11 , 12 , 18 لها منوال واحد فقط Unimodal هو 9 .

مثال: مجموعة البيانات التالية : 5 , 8 , 12 , 11 , 10 , 15 , 16 ليس لها منوال.

مثال: مجموعة البيانات التالية : 8 , 7 , 7 , 7 , 5 , 5 , 4 , 4 , 4 , 3 , 2 لها منوالان (ذات منوالين Bimodal) هما 7 , 4 .

المنوال في حالة البيانات المبوبة (الجداول التكرارية) :

في حالة البيانات المبوبة أو الجداول التكرارية لا نستطيع تحديد القيمة الأكثر تكراراً لأن القيم تكون مفهرسة داخل الفئات المختلفة ، وبذلك نستطيع أن نحدد الفئات المنوالية بأنها الفئات التي يقابلها أعلى تكرار ، وفي حالة تساوي تكرار الفئة السابقة للفئة المنوالية مع تكرار الفئة اللاحقة للفئة المنوالية فإنه يمكن حساب قيمة المنوال بمركز الفئة المنوالية أي في منتصفها ، وفي حالة عدم تساويهما في التكرار فإنه يمكن حساب المنوال كالتالي:

١) نوجد أكبر تكرار f وعليه يمكن إيجاد التكرار السابق له وهو f_1 والتكرار اللاحق له وهو f_2 .
٢) نأخذ بداية الفئة المنوالية ونرمز لها بالرمز A وهي الفئة التي يقابلها أعلى تكرار f.
٣) نحدد طول الفئة المنوالية L وهو يساوي الفرق بين بداية الفئة المنوالية وبداية الفئة التالية لها ويتم تطبيق القانون التالي :

$$Mod = A + \left(\frac{f - f_1}{2f - f_1 - f_2}\right) L$$

مثال: احسب المنوال لأعمار الطلاب المسجلة في الجدول التالي:

فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل:

من الجدول نحدد القيم التالية :

$f = 8,\quad f_1 = 5,\quad f_2 = 4,\quad A = 8.5$

$L = 10.5 - 8.5 = 2$

ثم نعوض في قانون المنوال :

$$Mod = 8.5 + \left(\frac{8-5}{16-5-4}\right) 2$$

$$Mod = 9.36$$

مثال: احسب المنوال لدرجات الحرارة في أحد الأشهر المسجلة في جدول التوزيع التكراري التالي:

درجات الحرارة	عدد الأيام
-30	2
-32	6
-34	11
-36	8
40-38	3
المجموع	**30**

الحل :

من الجدول نحدد القيم التالية :

$f = 11\quad f_1 = 6\quad f_2 = 8\quad A = 34$

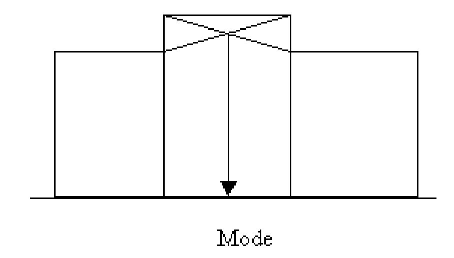
$L = 34 - 32 = 2$

ثم نعوض في قانون المنوال :

$$Mod = 34 + \left(\frac{11-6}{22-6-8}\right) \times 2$$

$$Mod = 35.25$$

كما يمكننا أن نجد المنوال بالرسم وذلك برسم المدرج التكراري لأكبر ثلاث فئات، ليكون المنوال كما هو في الشكل التالي :



<u>مميزات المنوال :</u>

١) مقياس سهل حسابه ولا يتأثر بالقيم الشاذة.
٢) يمكن إيجاده للقيم الوصفية والتوزيعات التكرارية المفتوحة.

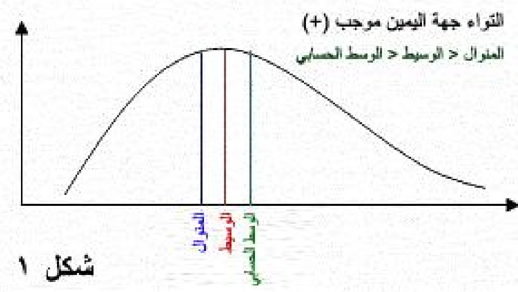
<u>عيوب المنوال :</u>

١) لا يأخذ جميع قيم البيانات في الاعتبار.
٢) قد يكون لبعض البيانات أكثر من منوال وبذلك لا يمكن تحديد قيمة وحيدة للمنوال.

العلاقة بين الوسط الحسابي والوسيط والمنوال (أشكال الالتواء Skewness):

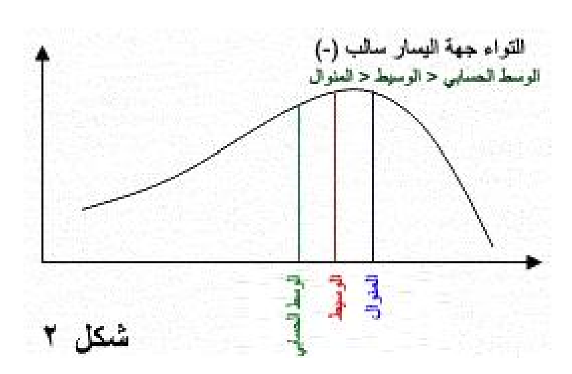
إن الوضع النسبي للوسط والوسيط والمنوال على المنحنيات التكرارية توصف بأربعة أشكال :

١) التواء موجب (Skewness): عندما يكون المنحنى غير متماثل أي مفرطح والتفرطح جهة اليمين (+) فإن: المنوال < الوسيط < الوسط الحسابي (شكل١).

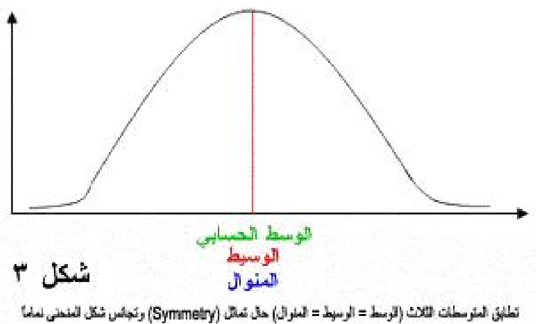


شكل ١

٢) **التواء سالب (Skewness): عندما يكون المنحنى غير متماثل أي مفرطح والتفرطح جهة اليسار (–) فإن: الوسط الحسابي< الوسيط < المنوال (شكل٢).**



٣) **حالة التماثل (Symmetry): تطابق المتوسطات الثلاث (الوسط = الوسيط = المنوال) حال تماثل وتجانس شكل المنحنى تماماً (شكل٣).**



٤) **تفرطح معتدل : منحنى التوزيع يكون التفرطح فيه معتدل ويكون:**

الوسط الحسابي – المنوال = ٣ (الوسط الحسابي – الوسيط).

خامساً : الوسط الهندسي (Geometric Mean)

يرمز له بالرمز G ، وفي بعض الكتب الأخرى يرمز له بالرمز G.M فالوسط الهندسي لمجموعة n من الأرقام $X_1, X_2, \ldots, X_n$ هو الجذر النوني لحاصل ضرب هذه الأرقام:

$$G = \sqrt[n]{X_1, X_2, X_3, \ldots, X_n}$$

ويمتاز الوسط الهندسي عن الوسط الحسابي بأنه أقل تأثراً بالقيم الشاذة في البيانات لأنه معلوم رياضياً بأن الوسط الهندسي لمجموعة من القيم أقل من وسطها الحسابي.

مثال: احسب الوسط الهندسي للأرقام 2 , 4 , 8 ؟

الحل :

$$G = \sqrt[3]{(2)(4)(8)} = \sqrt[3]{64} = 4$$

ويمكن حساب الوسط الهندسي باستخدام اللوغاريتمات كالتالي :

$$\text{Log G.M.} = \frac{1}{n}\left(\sum_{i=1}^{n} Log(x_i)\right)$$

مثال: احسب الوسط الحسابي والوسط الهندسي للأعداد : 3 , 5 , 6 , 6 , 7 , 10 , 12

الحل :

بحساب الوسط الهندسي باللوغاريتمات :

$$\text{G.M.} = \sqrt[7]{3.5.6.6.7.10.12}$$

$$\text{Log G.M.} = \frac{1}{7}(\text{Log } 3 + \text{Log } 5 + \text{Log } 6 + \text{Log } 6 + \text{Log } 7 + \text{Log } 10 + \text{Log } 12)$$

$$= \frac{1}{7}(0.4771 + 0.699 + 0.7782 + 0.7782 + 0.8451 + 1 + 1.0729)$$

$$= 0.8081$$

$$\text{G.M.} = 6.43$$

بحساب الوسط الحسابي:

$$\bar{x} = \frac{1}{7}(3 + 5 + 6 + 6 + 7 + 10 + 12) = 7$$

ونلاحظ أن قيمة الوسط الحسابي $\bar{x}$ من المثال أكبر من الوسط الهندسي .G.M وهذا يوضح الحقيقة بأن الوسط الهندسي لمجموعة أرقام موجبة غير متساوية أقل من وسطها الحسابي.

سادساً : **الوسط التوافقي** (Harmonic Mean)

يستخدم الوسط التوافقي عندما يكون مقلوب المتغير له دلالة كأن يعين نسبة بين متغيرين مرتبطين مثل السرعة بالنسبة للزمن ، ويرمز له بالرمز H ، فالوسط التوافقي لمجموعة من n من الأرقام $X_1, X_2, \ldots, X_n$ هو مقلوب الوسط الحسابي لمقلوبات هذه القيم .

$$H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}} = \frac{n}{\sum\frac{1}{X}}$$

ومن الناحية العملية فإنه من الأسهل أن نتذكر أن :

$$\frac{1}{H} = \frac{\sum\frac{1}{X}}{n} = \frac{1}{n}\sum\frac{1}{X}$$

مثال: احسب الوسط التوافقي للبيانات : 3 , 5 , 6 , 6 , 7 , 10 , 12

الحل :

$$\frac{1}{H} = \frac{1}{7}\left(\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{6}+\frac{1}{10}+\frac{1}{12}\right)$$

$$\frac{1}{H} = \frac{1}{7}\left(\frac{140+84+70+70+60+42+35}{420}\right)$$

$$\frac{1}{H} = \frac{501}{2940} \Rightarrow H = 5.87$$

عند مقارنة نتيجة هذا المثال بنتائج المثال السابق لقيمتي الوسط الهندسي والوسط الحسابي يتضح أن : *الوسط التوافقي > الوسط الهندسي > الوسط الحسابي*

$$H < G < \overline{X}$$

وتتحقق علامة التساوي إذا كانت الأرقام $X_1, X_2, \ldots, X_n$ متساوية .

سابعاً: الربيعات والعشيرات والمئينات (Quartiles, Deciles، Percentiles):

إذا رتبت مجموعة من الأرقام حسب قيمها تصاعدياً أو تنازلياً فإن القيمة التي في المنتصف والتي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد هي (الوسيط Median) . وبتعميم هذه الفكرة يمكن أن نفكر في القيم التي تقسم المجموعة إلى أربعة أجزاء متساوية . هذه القيم يرمز لها بالرموز Q_1, Q_2, Q_3 تسمى بالربيع الأول ، الربيع الثاني ، الربيع الثالث على الترتيب. ويجب ملاحظة أن القيمة Q_2 تساوي الوسيط . كذلك فإن القيم التي تقسم المجموعة إلى عشرة أجزاء متساوية تسمى بالعشيرات فيرمز لها بالرمز $D_1, D_2, D_3, \ldots, D_9$ بينما القيم التي تقسم البيانات إلى مائة قسم متساو تسمى بالمئينات ويرمز لها بالرمز $P_1, P_2, P_3, \ldots, P_{99}$. العشير الخامس والمئين الخمسون يساويان الوسيط كما أن المئين الخامس والعشرون والمئين الخامس والسبعون يساويان الربيع الأول والربيع الثالث على الترتيب . وإجمالاً يمكن إيجاد الربيعات والعشيرات والمئينات وغيرها من القيم بتقسيم البيانات إلى أقسام جزئية متساوية تسمى قيم التقسيمات الجزئية.

الباب الرابع

مقاييس التشتت MEASURES OF DISPERSION

مقاييس النزعة المركزية التي تم مناقشتها في الباب السابق لا تعتبر كافية لوصف مجموعة من البيانات وصفاً كاملاً فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات). فالعينات التالية لها نفس الوسط الحسابي (8) ولكنها بلا شك تختلف عن بعضها ؛ فالعينة الأولى مفرداتها متساوية بينما العينة الثانية يوجد فرق بين أصغر مفردة وأكبر مفردة ؛ ونجد الفرق أكبر في العينة الثالثة ؛ أي أن العينات الثلاث مختلفة التجانس رغم تساويها في الوسط الحسابي .

عينة 1	8	8	8	8	8
عينة 2	3	4	6	11	16
عينة 3	2	5	7	8	18

فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى التفاف أو بعثرة البيانات حول هذا الوسط ، ولهذا نشأت الحاجة إلى إيجاد مقاييس تقيس درجة (تجانس) أو تشتت (تباعد) مفردات البيانات عن بعضها البعض ، وتعرف هذه المقاييس بمقاييس التشتت ، وسوف نستعرض منها كلٍ من : المدى ، نصف المدى الربيعي ، الإنحراف المتوسط ، التباين والإنحراف المعياري.

أولاً:المدى ونصف المدى (Range and Midrange):

يعتبر المدى مقياساً للتشتت من السهل حسابه ويعطي فكرة سريعة عن طبيعة البيانات، ويستخدم عادة في وصف الأحوال الجوية ، ولحساب المدى لمجموعة ذات بيانات مباشرة نقوم بطرح أصغر رقم من أكبر رقم ، ويرمز له بالرمز R .

المدى R = أكبر قيمة ـ أصغر قيمة

أما في حالة البيانات المبوبة فإن المدى يعرف بأكثر من طريقة نذكر منها الطريقتين التاليتين :

١) المدى R = الفرق بين مركزي الفئة العليا والفئة الدنيا.
٢) المدى R = الحد الأعلى للفئة العليا مطروحاً منه الحد الأدنى للفئة الدنيا.

مثال: احسب المدى R للمفردات التالية : 80 , 30 , 70 , 62 , 40 , 82

الحل : أكبر قيمة = 82 ، أصغر قيمة = 30

$$R = 82 - 30 = 52$$

مثال: أوجد المدى R من البيانا في الجدول التالي :

90 - 99	80 - 89	70 - 79	60 - 69	50 - 59	40 - 49	الفئات
1	2	11	15	9	2	عدد الموظفين

الحل : يمكننا الحل بطريقتين :

الأولى : نحسب مركز الفئة العليا ويساوي 94.5 ونحسب مركز الفئة الدنيا ويساوي 44.5 ، وبذلك نحسب المدى R :

$$R = 94.5 - 44.5 = 50$$

الثانية : الحد الأعلى للفئة العليا الحقيقي يساوي 99.5 والحد الأدنى للفئة الدنيا يساوي 39.5 ، وبذلك يكون المدى R :

$$R = 99.5 - 39.5 = 60$$

نلاحظ الاختلاف في النتائج السابقة ، وبذلك نحبذ دائماً استخدام الطريقة الأولى لحساب المدى R .

ثانياً: نصف المدى الربيعي أو الإنحراف الربيعي (Interquartile Range):

من أهم عيوب المدى Range أنه يتأثر بالقيم الشاذة وبالتالي لا يعطي المدى صورة صادقة عن البيانات ؛ لذلك دعت الحاجة إلى إيجاد مقياس آخر وهو نصف المدى الربيعي أو الانحراف الربيعي (Interquartile Range) ويرمز له بالرمز Q ويحسب من العلاقة التالية :

$$Q = \frac{Q_3 - Q_1}{2}$$

نتبع الخطوات التالية لحساب نصف المدى الربيعي :

١) نرتب البيانات تصاعديا.
٢) نوجد قيمة Q_1 وهي القيمة التي يسبقها ربع البيانات.
٣) نوجد قيمة Q_3 وهي القيمة التي يسبقها ثلاثة أرباع البيانات.
٤) ثم نطبق العلاقة التالية : $Q = \frac{Q_3 - Q_1}{2}$

مثال: أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب التالية :

67,65,69,58,55,71,72,70

الحل : نرتب البيانات تصاعدياً : 55,58,65,67,96,70,71,72

$Q_1 = \frac{58+65}{2} = 61.5$, $Q_3 = \frac{70+71}{2} = 70.5$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{70.5 - 61.5}{2} = 4.5$$

مثال: أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب التالية :

59,67,65,69,58,55,70,72,74

الحل: نرتب البيانات تصاعدياً : 55,58,59,65,67,69,70,72,74

$Q_1 = 59$, $Q_3 = 70$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{70 - 59}{2} = 5.5$$

ثالثاً: الإنحراف المتوسط (Absolute Mean Deviation):

يعرف بأنه متوسط الإنحرافات المطلقة للبيانات عن وسطها الحسابي $\bar{X}$ ويرمز له بالرمز .M.D ويحسب رياضياً من العلاقة التالية :

$$M.D = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n} = \frac{\sum |X - \bar{X}|}{n}$$

حيث n مجموعة الأرقام $X_1, X_2, \ldots, X_n$ و $\bar{X}$ هو الوسط الحسابي للأرقام و $|X_i - \bar{X}|$ هو القيم المطلقة لانحرافات القيم X_i عن $\bar{X}$.

مثال: أوجد متوسط الإنحرافات لمجموعة الأرقام : 2,3,6,8,11

الحل :

$$\bar{X} = \frac{2+3+6+8+11}{5} = 6$$

$$M.D = \frac{|2-6| + |3-6| + |6-6| + |8-6| + |11-6|}{5}$$

$$= \frac{|-4| + |-3| + |0| + |2| + |5|}{5}$$

$$= \frac{4+3+0+2+5}{}$$

$$= \frac{14}{5} = 2.8$$

أما إذا كانت $X_1, X_2, \ldots, X_k$ تحدث بتكرارات $f_1, f_2, \ldots, f_k$ على الترتيب فإن الإنحراف المتوسط يمكن كتابته على الصورة :

$$M.D = \frac{\sum_{i=1}^{n} f|X_i - \bar{X}|}{n} = \frac{\sum f|X - \bar{X}|}{n}$$

رابعاً:التباين والإنحراف المعياري (Variance and Standard Deviation):

التبـاين (Variance) : إذا أعطيت مجتمـع محـدود $X_1,X_2,...,X_n$ فـإن تبـاين المجتمـع يرمز له بالحرف σ^2 ويقرأ (مربع سيجما) :

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}$$

أما إذا كانت مشاهدات المجتمع تتكرر بتكرارات $f_1,f_2,...,f_k$ فإن التباين يحسـب مـن العلاقة التالية:

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i(X_i-\mu)^2}{\Sigma f_i} = \frac{\Sigma f_i(X_i-\mu)^2}{N}$$

الإنحـراف المعيـاري (Standard Deviation) : نحسب الإنحـراف المعيـاري بأخذ الجذر التربيعي للتبـاين ، ويرمـز للإنحـراف المعيـاري للمجتمـع بـالرمز σ بينمـا يرمـز للإنحـراف المعيـاري لعينـة مسحوبة من هذا المجتمـع بـالرمز S ، ونستخدم العلاقة التالية لحساب الإنحراف المعياري لمجموعة n للمشاهدات $X_1,X_2,...,X_n$:

$$S = \sqrt{\sum_{i=1}^{n}\frac{(X_i-\overline{X})^2}{n-1}} = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n-1}} = \sqrt{\frac{\Sigma X_i^2}{n-1}}$$

حيث أن X_i تمثل انحرافات كل رقم X_i عن $\overline{X}$ ، وعلى هذا فان S هي الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي ويسمى أحياناً جذر متوسط مربع الانحراف .

إذا كانت $X_1,X_2,...,X_k$ تحدث بتكرارات $f_1,f_2,...,f_k$ على الترتيب فان الانحراف المعياري يمكن كتابته على صورة :

$$S = \sqrt{\sum_{i=1}^{k}\frac{f_i(X_i-\overline{X})^2}{n-1}} = \sqrt{\frac{\Sigma f(X-\overline{X})^2}{n-1}} = \sqrt{\frac{\Sigma fx^2}{n-1}}$$

طريقة سريعة لتقدير الانحراف المعياري: إذا أردنا الحصول على قيمة تقديرية للانحراف المعياري فإننا غالباً ما نأخذ القيمة التالية:

$$S = \frac{Range}{4} = \frac{\text{المدى}}{4}$$

مثال: أوجد الإنحراف المعياري والتباين لأوزان ١٠٠ طالب جامعة المدرجة بيناتهم في الجدول التالي :

فئات الأوزان	60-62	63-65	66-68	69-71	72-74
عدد الطلاب	5	18	42	27	8

الحل: لابد من كتابة الجدول بالصورة التالية :

$f(x-\bar{x})^2$	$(x-\bar{x})^2$	$x-\bar{x}$	التكرار×مركز الفئة fx	التكرار f	مركز الفئة x	الفئات
208.0125	41.6025	-6.45	305	5	61	60-62
214.2450	11.9025	-3.45	1152	18	64	63-65
8.5050	0.2025	-0.45	2814	42	67	66-68
175.5675	6.5025	2.55	1890	27	70	69-71
246.4200	30.8025	5.55	584	8	73	72-74
852.7500			6745	100		المجموع

$$\overline{X} = \frac{\sum_{i=1} fX}{n} = \frac{\Sigma fX}{n} = \frac{6745}{100} = 67.45$$

$$S = \sqrt{\frac{\Sigma f(X-\overline{X})^2}{n-1}} = \sqrt{\frac{852.7500}{100-1}} = \sqrt{8.6136} = 2.93$$

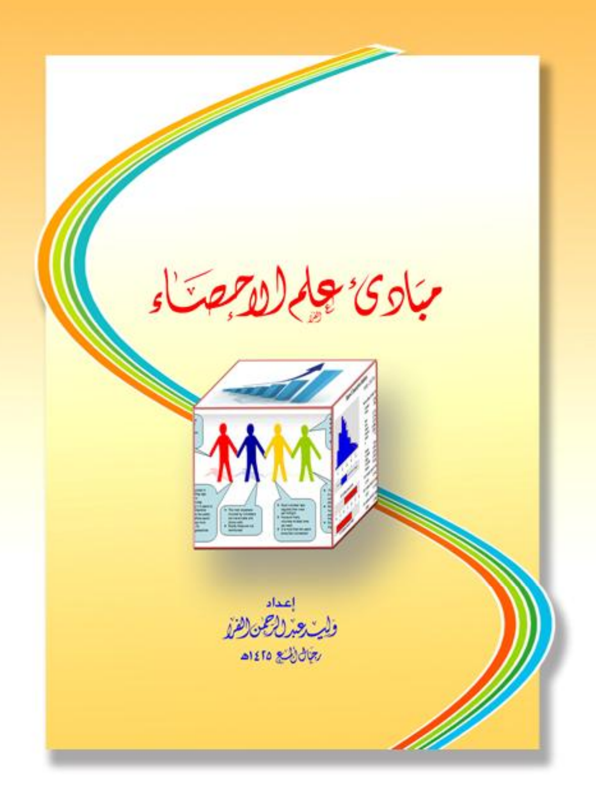
$$S^2 = 8.6$$

" تم بحمد الله "

المراجع

REFERENCES

١) أحمد عباده سرحان ، (١٩٦٨م) ، مقدمة في طرق التحليل الإحصائي ،معهد الدراسات والبحوث الإحصائية ـ جامعة القاهرة.
٢) جلال الصياد ومحمد الدسوقي حبيب ، (١٩٩٠م)، مقدمة في الطرق الإحصائية –الطبعة الثانية- تهامة – جدة – المملكة العربية السعودية.
٣) أنيس إسماعيل كنجو ، (١٩٩٣م)، الإحصاء والإحتمال – جامعة الملك سعود- عمادة شئون الطلاب.
٤) محمد صبحي أبو صالح وعدنان محمد عوض ،(١٩٨٣م) ، مقدمة في الإحصاء –الطبعة الرابعة – دار جون وايلي وأبنائه – نيويورك.
٥) سمير كامل عاشور وسامية سالم أبو الفتوح ، (١٩٩٠م)، مقدمة في الإحصاء الوصفي - معهد الدراسات والبحوث الإحصائية ـ جامعة القاهرة.



محافظة رجال ألمع 1425 هـ _ منطقة عسير _ المملكة العربية السعودية _ وليد الفرا - E.Mail: walfarra@hotmail.com